Table 6.2 DAMAGE FUNCTIONS

Constituent name	Units	Level of damage: None 0	Excellent 2	Acceptable 4	Slightly polluted 6	Polluted 8	Heavily polluted 10	Reference*
Aluminum	mg/l	0	0.01	0.05	0.10	.50	1.00	34
Ammonia	mg/l	0	0.1	0.3	0.9	2.7	3.0	27
Dissolved oxygen	mg/l	>9	8.0	6.8	4.5	1.8	0.9	32
Inorganic carbon	mg/l	<50	70	90	110	130	150	32
Chloride	mg/l	0	25	175	200	240	250	28
Chloroform extract	mg/l	0	0.04	0.15	0.25	0.35	0.40	28
Chromium	mg/l	0	0.02	0.05	1.0	10.0	50.0	33,34
Coliforms-total	MPN/100ml	0	100	2000	7500	15,000	150,000	28,33
Coliforms-fecal	MPN/100ml	0	20	200	800	3,000	50,000	29,32
Copper	mg/l	0	0.02	0.10	1.00	5.00	10.00	33,34
Cyanide	mg/l	0	0.01	0.02	0.05	0.10	0.50	33,34
Fluoride	mg/l	<0.7	0.8	0.9	1.2	3.0	8.0	34
Iron	mg/l	0	0.1	0.3	0.9	2.7	3.0	27
Lead	µg/l	0	5	50	100	250	350	33,34
Manganese	mg/l	0	0.05	0.17	0.50	1.00	1.50	27
Mercury	µg/l	0	1	5	10	20	50	34
Nickel	mg/l	0	0.01	1.0	3.0	9.0	20.0	34
Inorganic nitrogen	mg/l	<0.6	0.9	3.0	4.5	7.0	10.0	32
Oil-grease	mg/l	0	0.01	0.10	5	30	50	34
pH-MIN		7	6.5	6.0	5.0	4.0	3.9	27
pH-MAX		7	8.0	8.4	9.0	10.0	10.1	27
Phenol	µg/l	0	0.5	1.0	20	100	200	33,34
Phosphates	mg/l	0	0.1	0.2	0.5	1.6	10	29
Solids-dissolved	mg/l	<100	200	500	1000	1500	2300	32
Solids-suspended	mg/l	0	20	40	100	280	300	27
Temp. diff.	°C	0	1.0	2.5	3.0	4.0	10.0	29
Tin	mg/l	0	10	40	100	300	1000	33,34
Zinc	mg/l	0	0.1	1	5	15	40	34

*The references shown are those used to develop the damage function for each constituent.

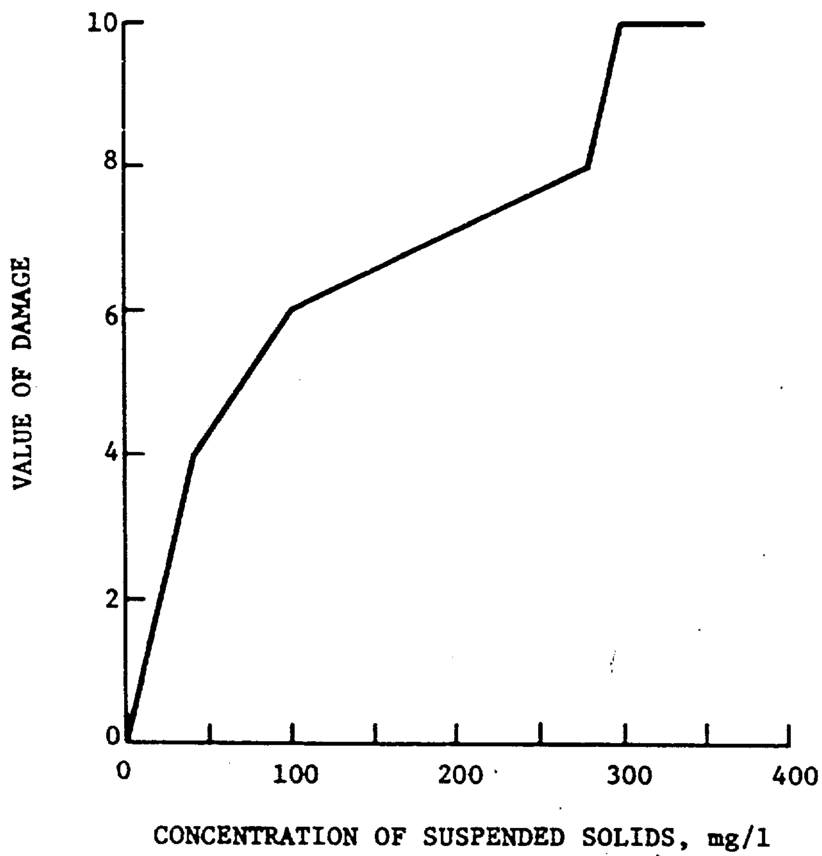

Figure 6.3 Example damage function

It should be noted that the damage functions given in Table 6.2 exhibit just one of many possible choices for damage functions definitions. A monitoring agency should feel free to modify or change the damage functions as it sees fit, especially as more detailed and extensive reports relating damages to water quality become available. Also, the damage functions could be completely eliminated, if desired, by setting them equal to one for all values of the water quality indicators. This would eliminate damage in the resource allocation criterion, the choice of sampling frequency would then just depend on the probability of violation.

VI.3 FORMULATION OF "COST" OF UNDETECTED VIOLATIONS

The "cost" of undetected violations will now be derived. For the present it will be assumed that only one set of effluent standards is given for each source. This corresponds to the case in which there is a single outfall or the permits are written for the combined discharge from several outfalls. The case where there are several sets of standards for several outfalls will be treated at the end of this section.

Let M_{ij} be the mass loading of the j^{th} constituent from the i^{th} source. M_{ij} is modeled as either a normal or lognormal random variable with known mean and standard deviation. Let ϕ_{ij} be the density function of M_{ij}. The concentration of the corresponding stream parameter below the source is

$$CO_{ij} = a_{ij}M_{ij} + b_{ij} \tag{6.10}$$

where a_{ij} and b_{ij} depend on the effluent and upstream flow, the assumed upstream concentration, and, where needed, other stream parameters* (see Section VI.1 and Appendix C). All the quantities needed to calculate a_{ij} and b_{ij} in (6.10) are readily available to the monitoring

---

* If the effluent standard is written in terms of the concentration of pollutant, the formula for CO_{ij} is in the same form as (6.10) with M_{ij} replaced by CS_{ij} QS_{ij} where CS_{ij} is the concentration of pollutant j in the effluent and QS_{ij} is the source flow.

agency except for the concentration of the water quality indicators upstream from the source (denoted CU_{ij}). Even if knowledge of CU_{ij} were available, it does not make sense to use an actual value of upstream concentration in the priority procedure. This can be seen by considering the case of two similar plants, one slightly upstream from the other on the same river. If the actual upstream concentration were used, the plant further downstream would always be causing more damage (as measured by the downstream concentration of pollutants); this clearly is not equitable. If only the change in damage were considered, then the plant upstream would typically be penalized, since the change in damage for a given increase in pollutant concentration typically gets smaller as the concentration of pollutant increases. In other words, most of the damage functions are concave in shape. Instead of using an actual value of upstream concentration, it is suggested that the upstream concentration of all the pollutants in the basin be set so as to cause the same level of damage immediately upstream from each source. Clearly, the concentrations could be set to cause no damage upstream; this corresponds to setting the upstream concentration to zero for most water quality indicators. Nonzero initial damages might be desired in order to eliminate the sensitivity of the priority procedure to the damage function definitions for small values of damage. This method of setting CU_{ij} is equitable and is consistent with the priority procedure.

The damage due to the j^{th} constituent from the i^{th} source is $D_j(CO_{ij})$, where D_j is the damage function for the j^{th} constituent. Note that $D_j(CO_{ij})$ is a random variable with statistics depending on the statistics of M_{ij}. The expected damage due to the j^{th} constituent of the i^{th} source is then

$$D_{ij} = E\{D_j(CO_{ij})\} \tag{6.11}$$

or (using (6.10))

$$D_{ij} = \int D_j({}_{ij}M + b_{ij})\ \phi_{ij}(M)dM \tag{6.12}$$

The calculation of D_{ij} is carried out in detail in Appendix C. The expected damage from all the constituents of the i^{th} source (if unmonitored) is

$$c_i = \max_j D_{ij} \tag{6.13}$$

since the pollutant that causes the most damage is the one that limits the water quality. Note that (6.13) assumes that there is no synergistic or antagonistic interaction among the pollutants. This assumption is valid in general. For the purposes of this report, the extra complexity needed if this assumption were to be dropped is not warranted. The total damage that can be expected from all the sources, n_s, in the region is then

$$\sum_{i=1}^{n_s} c_i. \tag{6.14}$$

Taking the damage as additive corresponds to assuming noninteraction between the various sources.

The derivation leading to (6.14) did.not take into account the fact that we are only interested in undetected violations. The effect of the monitoring on the "cost" will be accounted for as follows: it is assumed that if, during the period of consideration, one of the constituents of a given source is found to exceed its standard, say τ_{ij}, the purpose of the monitoring has been achieved and the "cost" due to that source will be considered zero. Consequently, a violation is declared if at least one constituent is found to exceed its specified standard. Let p_{ij} be the probability that <u>no violation</u> will be observed in one sample of M_{ij}, i.e.,

$$p_{ij} = \int_0^{\tau_{ij}} \phi_{ij}(M)dM , \tag{6.15}$$

In view of the above discussion, the expected <u>"cost" of undetected violations</u> is obtained as follows. Assume that the i^{th} effluent is sampled

s_i times during the period consisting of N intervals (e.g., days). Denote by V_i the event that a violation is observed when sampling the effluent. Then, the total "cost" incurred due to the i^{th} effluent when it is sampled s_i times is, using the total probability law, the expected damage given that the standard violation was not detected times the probability that the violation was not observed plus the expected damage given that a violation was detected times the probability that a violation will be observed. Mathematically this can be written

$$C_i(s_i) = \frac{1}{N} E \left\{ \sum_{k=1}^{N} d_i(k) | \overline{V}_i \right\} P \left\{ \overline{V}_i | s_i \right\}$$

$$+ \frac{1}{N} E \left\{ \sum_{k=1}^{N} d_i(k) | V_i \right\} P \left\{ V_i | s_i \right\} \tag{6.16}$$

where $d_i(k)$ is the damage incurred due to the i^{th} source during the k^{th} interval and where $\overline{V}_i$ denotes the event that no violation is observed when sampling the source. The division by N, the number of intervals in the monitoring period, is just a normalization factor so that the damage is averaged over the monitoring period. If a violation is detected, the cost is zero, i.e., the second term on the right hand side above is zero. This follows from the fact that we are dealing with the cost of <u>undetected</u> violations and a <u>detected</u> violation should not enter in this cost. Therefore (6.16) becomes

$$C_i(s_i) = \frac{1}{N} E \left\{ \sum_{k=1}^{N} d_i(k) | \overline{V}_i \right\} P \left\{ \overline{V}_i | s_i \right\}. \tag{6.17}$$

Dropping the time dependence (variable k), one has

$$C_i(s_i) = c_i \; P(\overline{V}_i | s_i) \tag{6.18}$$

where c_i, the expected damage due to source i, is given in (6.13). The probability that no violation is observed when the source is sampled s_i times is (assuming independence between the concentrations at various sampling times)

$$P\{\overline{V}_i | s_i\} = p_i^{s_i} \tag{6.19}$$

Substituting this into (6.18) one obtains

$$C_i(s_i) = c_i \; p_i^{s_i} \tag{6.20}$$

The calculation of p_i, the probability that source i will not be in violation in one sampling, depends on the probability each constituent of source i will not be in volation, p_{ij} , and on the statistical.dependence between the various constituents.
observed in <u>one</u> sampling of source i, assuming independence between the various constituents, is

$$p_i = \prod_j p_{ij} \tag{6.21}$$

If the constituents are completely correlated, then

$$p_i = \min_j p_{ij} \tag{6.22}$$

Since data are not readily available to ascertain the exact correlation between the various constituents of a source, either complete dependence or independence must be assumed.

The "cost" of undetected violation is, therefore,

$$C = \sum_{i=1}^{n_s} C_i(s_i)$$

$$= \sum_{i=1}^{n_s} c_i p_i^{s_i} \tag{6.23}$$

where, for the i^{th} source, s_i is the number of times the source is monitored, c_i is the expected damage, and p_i is the probability the source will not be found in violation if it is monitored once.

It remains to consider the case where there are several outfalls, each with its own set of standards. The outfalls can flow into one stream or into different streams. First consider the case where the outfalls flow into a single stream. The damage depends on the total mass load of pollutants. Assuming the outfalls lie close to each other, the expected damage can then be calculated in the usual way, using a combined mass load and flow rate. This is discussed further in Appendix C. Let $D_{ij\ell}$ be the expected damage due to the j^{th} constituent from source i into stream ℓ. The expected damage due to the i^{th} source from all the constituents into all streams is then (analogous to (6.13))

$$c_i = \max_{j,\ell} D_{ij\ell} \tag{6.24}$$

The calculation of the probability of no violation is straightforward since, assuming that the effluents from the various outfalls are independent, the probability of no violation from all the outfalls is the product of the probability of no violation in each of the outfalls. To be precise, let

p_{ijk} = probability of no violation due to pollutant j, outfall k, source i.

p_{ijk} is calculated analogously to (6.15). Using (6.21) and (6.22), the probability of no violation of any standard from outfall k, source i is

$$p_{ik} = \begin{cases} \Pi\, p_{ijk} & ; \quad \text{uncorrelated constituents} \\ \min p_{ijk} & ; \quad \text{correlated constituents} \end{cases} \tag{6.25}$$

The probability of no violation from any pollutant of any outfall for the source i is then

$$P_i = \prod_k P_{ik} \tag{6.26}$$

where we have assumed that the pollutant loadings in the outfalls are independent. Expected damage and probability of violation have been calculated for a source having many outfalls. The "cost" of undetected violations for this source can then be calculated using (6.20).

Example

In this subsection the "cost" of undetected violations is calculated for a simple case. Consider a single source having two constituents: suspended solids and BOD_5. The various parameters needed to calculate the "cost" are given in Table 6.3. Figures 6.4 and 6.5 give, respectively, the density functions for suspended solids and BOD_5. The probability of violating the effluent standard is the area under the density curve in the region to the right of the effluent standard. This area is shaded in the figures. For this example, the probability of violating the standards for suspended solids is 26% and for BOD_5, 12%.

The relation between the downstream concentration CO (in mg/1) and the example parameters is of the form

$$CO = aM + b \tag{6.27}$$

where M is the mass loading. The formulas for a and b are given in Appendix C. For suspended solids

$$a = \frac{1}{QU + QS} = 0.83$$

and

$$b = CU\left(\frac{QU}{QU + QS}\right) = 0$$

and for dissolved oxygen

$$a = -K_{BOD-DO}/(QU + QS) = -0.42$$

Table 6.3 EXAMPLE PARAMETERS

Parameter	Suspended solids	BOD_5
Upstream flow - QU	1.0 Ml/day	1.0 Ml/day
Effluent flow - QS	0.2 Ml/day	0.2 Ml/day
Distribution	Lognormal	Normal
Mean of Loading - μ	1.5 log kg	1.5 kg
Stan. Dev. of Loading - σ	0.3 log kg	0.5 kg
Effluent Standard - τ	50 kg	2.5 kg
DO concentration of effluent - CS	-	4 mg/l
BOD_5-DO transfer coefficient - $K_{BOD\text{-}DO}$	-	0.5
Assumed upstream concentration - CU	0	0
Assumed upstream concentration of DO	-	9 mg/l

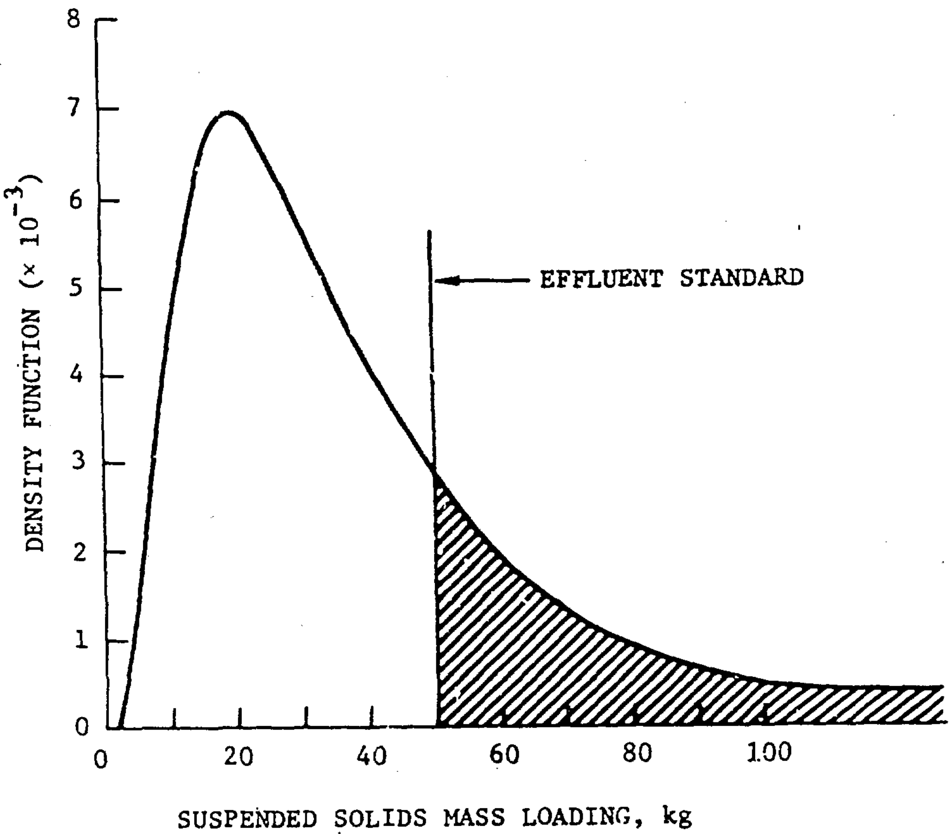

Figure 6.4 Example of density function for suspended solids

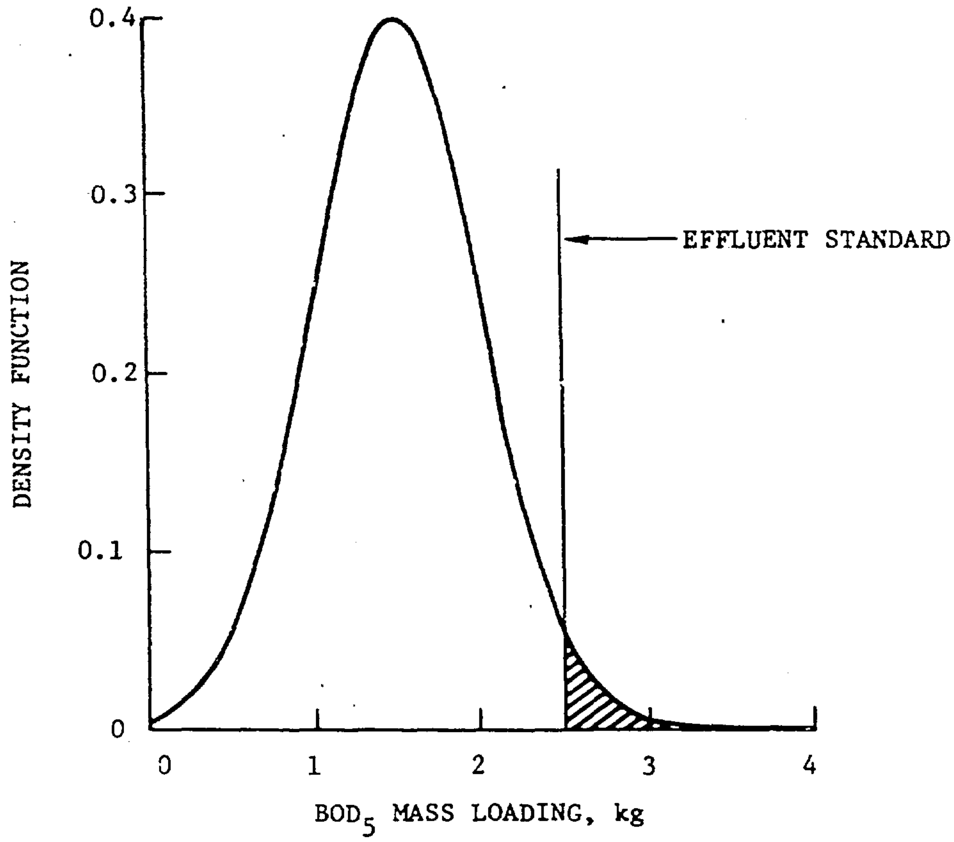

Figure 6.5 Example of density function for BOD_5.

and

$$b = \frac{1}{QU + QS} (CS_{DO}QS + CU_{DO}QU) - K_{BOD-DO}\ CU_{BOD}$$

$$= 8.2$$

Figure 6.6 shows the density function for CO, the concentration of suspended solids downstream from the source. Also shown is the damage function for suspended solids (note that the ordinate of the density function is not shown). The expected damage is just the area under the product of the density and damage function curves (see (6.12)). For this case the expected damage is 2.86. Figure 6.7 similarly shows the density and damage functions for dissolved oxygen. The expected damage resulting is 1.33. Therefore, we have the expected damage and probability of violation for these two parameters. Assume that the daily variations of the parameters are independent; then the probability, p_i, that the source will not be in violation is the product of the probabilities that each parameter will not be in violation (see (6.21)), or

$$p_i = (1-0.26) \times (1-0.88)$$
$$= 0.65$$

The expected damage from the source, c_i, is the maximum of the damages due to the individual constituents (6.13), so

$$c_i = \max \{2.86, 1.33\}$$

$$= 2.86$$

The "cost" of undetected violations for source i, given that the source was sampled s_i times, is

$$c_i(s_i) = c_i p_i^{s_i} \qquad (6.28)$$

Table 6.4 shows how the "cost" decreases, for this example, as the number of visits, s_i, increases.

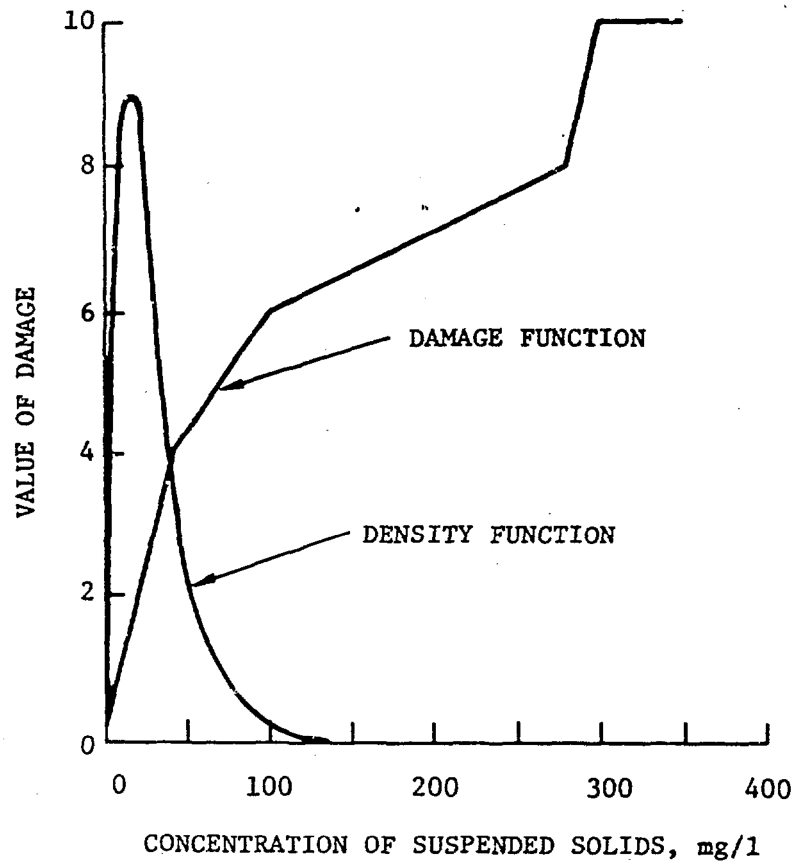

Figure 6.6 Density function and damage function for concentration of suspended solids in stream.

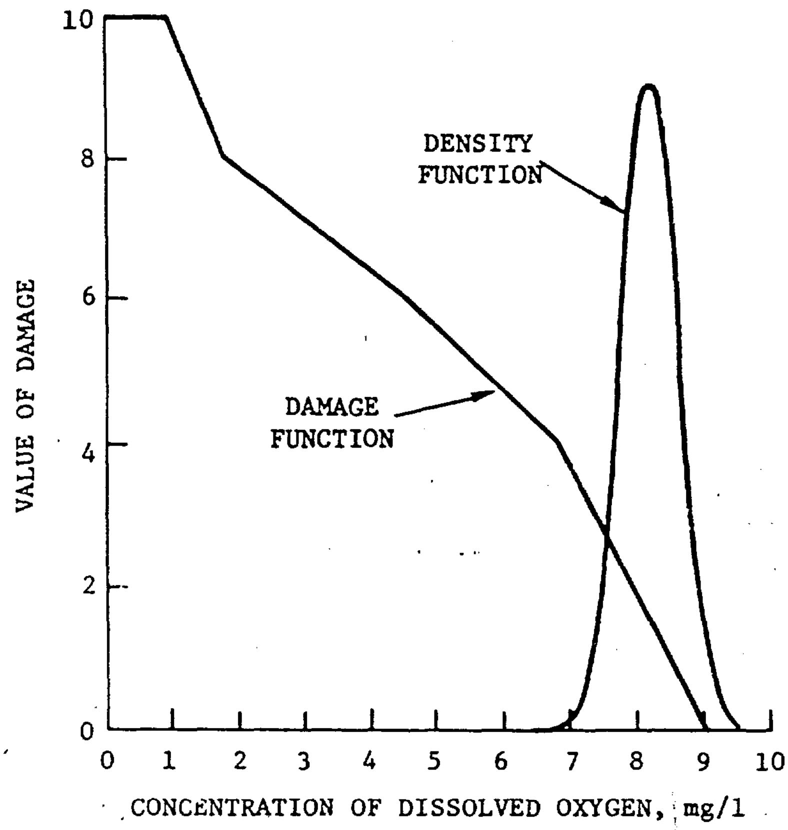

Figure 6.7 Density function and damage function for concentration of dissolved oxygen in stream.

Table 6.4 "COST" VERSUS NUMBER OF SAMPLES FOR EXAMPLE

s_i	"Cost" of undetected violations
0	2.86
1	1.86
2	1.23
3	0.80
4	0.52
5	0.34

SECTION VII
RESOURCE ALLOCATION PROBLEM

In the previous section, a performance criterion for the procedure of allocating monitoring resources was defined. This section defines the complete resource allocation problem and describes a method of solution - maximum marginal return.

VII.1 FORMULATION OF PROBLEM

There are three resource allocation problems that the monitoring agency might want solved:

1) Given a certain amount of resources (i.e. budget), determine how the monitoring resources should be allocated to minimize the "cost" of undetected violations.
2) In setting up a monitoring program, determine what level of resources is needed to insure that the "cost" of undetected violations is below a given level.
3) Given an increment of resources, determine how to allocate these additional resources and the resulting improvement in the monitoring system performance.

In the remainder of this subsection, these problems are formulated mathematically.

The "cost" of undetected violations (from Section VI.3) is

$$C(\underline{s}) = \sum_{i=1}^{n_s} C_i(s_i) \tag{7.1}$$

where $\underline{s} = (s_1, s_2, \ldots, s_{n_s})$,

$$C_i(s_i) = c_i p_i^{s_i} \tag{7.2}$$

c_i is the expected damage for the i th source, p_i is the probability no violation is observed at the i^{th} source, n_s is the number of sources, and s_i is the number of times the i^{th} source is monitored. The total cost to monitor all the sources where the i^{th} source is monitored s_i times is

$$R(s) = \sum_{i=1}^{n_s} r_i s_i \tag{7.3}$$

where r_i is the cost of monitoring source i once. r_i is made up of manpower, transportation, equipment and laboratory costs. The actual values of these costs will vary from agency to agency and as a function of time; they are therefore not specified in this report. r_i, however, is calculated for the demonstration case; see Section IX and Appendix D.

Upper and lower bounds on s_i may also be given, i.e.

$$\ell_i \leq s_i \leq L_i \tag{7.4}$$

To see where a monitor may desire to do this, consider the case where, from ambient monitoring, it has been observed that in a certain reach the level of a particular constituent is higher than usual. Then, one might want to check at least once during the next period all the effluent sources that might have caused this. In this case a lower bound of unity is set on the corresponding sampling rates. Also, consider the case of an effluent having a small expected violation cost. Based upon the existing information, it will have a low priority for being monitored. In order to prevent information from becoming obsolete, one can stipulate that it has to be monitored at least once during a certain period of time. An upper bound might be desired if the monitor does not want to sample any source more than a given number of times. This would be true, for example, if the monitor were required to visit a certain number of sources. Another situation can occur when there is a known polluter (e.g., one against which there are sufficient data to initiate legal action or one which is improving its treatment according to an approved long-term plan); the monitor may then decide not to survey this source frequently because the result is predictable. In this case, the upper bound for s_i would be set to some specified value.

The three optimization problems can now be specified. Problem 1 is

$$\begin{aligned} &\text{minimize } C(\underline{s}) \\ &\text{subject to } R(\underline{s}) \le B \qquad \text{(P1)} \\ &\underline{\ell} \le \underline{s} \le \underline{L} \end{aligned}$$

where B is the monitoring agency's budget and $\underline{\ell} = (\ell_1,\ldots,\ell_{n_s})$ and $\underline{L} = (L_1,\ldots,L_{n_s})$ are upper and lower bounds. Problem 2 is

$$\begin{aligned} &\text{minimize } R(\underline{s}) \\ &\text{subject to } C(\underline{s}) \le A \qquad \text{(P3)} \\ &\underline{\ell} \le \underline{s} \le \underline{L} \end{aligned}$$

where A is the maximum "cost" of undetected violations allowed. Problem 3 is of the same form as Problem 1, except B includes the additional resources and $\underline{\ell}$ specifies the sampling frequencies under the original allocation. The decrease in "cost" between when the original budget is used and the new budget is used is the system improvement. The additional samples specify where to use the additional resources.

VII.2 METHOD OF MAXIMUM MARGINAL RETURN

The optimization method used to solve the resource allocation problems is the method of maximum marginal return. It is particularly suited for these problems since it solves all three problems in the same manner.

The features of the method of maximum marginal return are:

(1) It is very fast on the computer. The computation time grows only proportionally with the size of the problem.

(2) If the function to be minimized is convex, this method will yield the absolute minimum when the cost of resource quanta is equal.

The cost C(s) can be easily shown to be convex--its second derivative is strictly positive for $s_i < N$ (which is always the case) and $p_i < 1$ (this is also satisfied, since p_i is a probability). The only condition that is not satisfied for Problem 1 is the requirement that cost of the quanta, r_i, be equal. However, the method will yield nearly the optimum allocation if

$$\max r_i \ll B \tag{7.5}$$

i.e., the largest cost of a sample is much smaller than the total budget B. Then the difference between the solution obtained by this method and the absolute minimum is negligible. Since (7.5) will be satisfied for the monitoring resource allocation problem, the maximum marginal return method is well suited for determining the sampling rates.

The method of maximum marginal return is basically a steepest descent algorithm. It is based on the following intuitive idea: the best place to allocate one unit of resource is where the marginal return (the decrease in cost - in our case undetected violation "cost" - accrued by using that unit of resource) is greatest. Therefore, by ordering the marginal returns in descending order, one obtains a priority list with the items having highest priority on top.

To be precise, the marginal return accrued when the sampling time on the i^{th} source is increased from s_i-1 to s_i is

$$\mu_i(s_i) = \frac{C_i(s_i-1) - C_i(s_i)}{r_i} \tag{7.6}$$

In view of the convexity of C_i, these marginal returns are monotonically decreasing with s_i, i.e.,

$$\mu_i(s_i) > \mu_i(s_i + 1) \tag{7.7}$$

The priorities of allocation are obtained by simply ordering these marginal returns. If the ordering obtained is, for example,

$$\mu_2(1) > \mu_1(1) > \mu_2(2) > \mu_3(1) \quad \ldots. \tag{7.8}$$

then effluent 2 is sampled with highest priority, then effluent 1, then again effluent 2, then effluent 3, etc. Following this, a relation between the minimized "cost" of undetected violations and the corresponding resource cost is obtained. Therefore, this method solves simultaneously the problem of minimizing the undetected violation "cost" subject to the total budget and the minimization of the budget subject to a given "cost" of undetected violations.

The problem of allocating an increment of resources to maximize the improvement in an existing monitoring system is solved as follows: Set up the priority list as described above, and remove from the list those samples that have been allocated. The remaining items on the list are, in descending priority, the ones that should be monitored with an increase in resources.

The above method will be illustrated via a simple numerical example. Assume there are $n_s=3$ pollutant sources with "costs" of undetected violation (for the period in consideration) as given in Table 7.1. It is assumed, for the purpose of this example, that costs of monitoring each of these effluents are the same (equal to one).

The "costs" C_i of undetected violations are given in these tables as functions of the corresponding number of samples. The maximum number of samples per source is taken as 5. Also the marginal returns as defined in (7.6) and the priority ordering according to (7.8) appear next to each sample.

The priority list of the sources sampled appears in Table 7.2 together with the "cost" of undetected violations as a function of the available resources. This table shows immediately the necessary resources to achieve a given "cost" of undetected violations and also the achievable minimum "cost" of undetected violations for a given amount of resources (number of samples).

As an example of Problem 1, consider the problem of finding the best allocation of 6 samples. From column 2 of Table 7.2, one sees that the 6 samples should be taken from sources 3, 2, 1, 3, 1, and 2. The sampling frequencies are then s_1 = 2, s_2 = 2, s_3 = 2. From column 3 of Table 7.2, the "cost" corresponding to these frequencies is 1.16.

As an example of Problem 2, consider the problem of finding the minimum amount of resources required to bring the "cost" of undetected violations to 1.00 or less. From column 3 of Table 7.2, one sees that the first time that the "cost" drops below 1.00 occurs for 7 samples, for

Table 7.1 "COST" OF UNDETECTED VIOLATIONS AND PRIORITY ORDERING

Number of samples	"Cost" undetected violations	Marginal return	Priority order
s_1	c_1	μ_1	
0	1.00		
1	0.70	0.30	3
2	0.45	0.25	5
3	0.25	0.20	7
4	0.08	0.17	8
5	0.02	0.06	14
s_2	c_2	μ_2	
0	1.00		
1	0.65	0.35	2
2	0.42	0.23	6
3	0.27	0.15	9
4	0.15	0.12	11
5	0.08	0.07	13
s_3	c_3	μ_3	
0	1.00		
1	0.55	0.45	1
2	0.29	0.26	4
3	0.15	0.14	10
4	0.05	0.10	12
5	0.01	0.04	15

which the cost is 0.96. From column 2 of this table one sees that the 7 samples should be taken, in order, from sources 3, 2, 1, 3, 1, 2, and 1. The corresponding sampling frequencies are thus $s_1 = 3$ (three samples at source 1), $s_2 = 2$ (two samples at source 2), and $s_3 = 2$ (two samples at source 3).

As an illustration of how to use the information to allocate additional resources to improve an existing monitoring system (Problem 3), assume that the preassigned sampling frequencies are

$$s_1 = 1, \; s_2 = 2, \; s_3 = 1$$

Consider the problem of optimally allocating four more samples. This is solved as follows: Take the priority list and omit the first s_i samples on source i, as illustrated in Table 7.3. Then it is seen that the priorities for the additional four samples are: first source #3, then #1, again #1, and again #1. The resulting overall sampling frequencies are

$$s_1 = 4, \; s_2 = 2, \; s_3 = 2.$$

Table 7.3 ALLOCATION OF ADDITIONAL INCREMENTS OF RESOURCES TO A GIVEN MONITORING SYSTEM

Original priority list of sources	Priority list of sources given the preassigned samples
3	
2	
1	
3	3
1	1
2	
1	1
1	1
2	2
3	3
2	2
3	3
2	2
1	1
3	3

SECTION VIII
RESOURCE ALLOCATION PROGRAM

Components of the allocation procedure were described in the previous three sections. This section discusses how these components fit together to form the Resource Allocation Program. Examples are also given showing the operation of the Program.

VIII.1 GENERAL PROGRAM DESCRIPTION

A flowchart of the Resource Allocation Program is shown in Figure 8.1. The following is a brief description of the function on the various components.

(1) Initialize Statistical Description

Combine the raw self-monitoring and compliance monitoring data to obtain an initial statistical description (distribution, mean and standard deviation) for each pollutant of each source.

(2) Calculate Expected Damage and Probability of Violation

Use the statistical description of the effluent loads, the effluent standards, and the stream parameters to obtain the expected damage and probability of violation for each source.

(3) Determine Priorities

Use the method of maximum marginal return to obtain the monitoring frequencies.

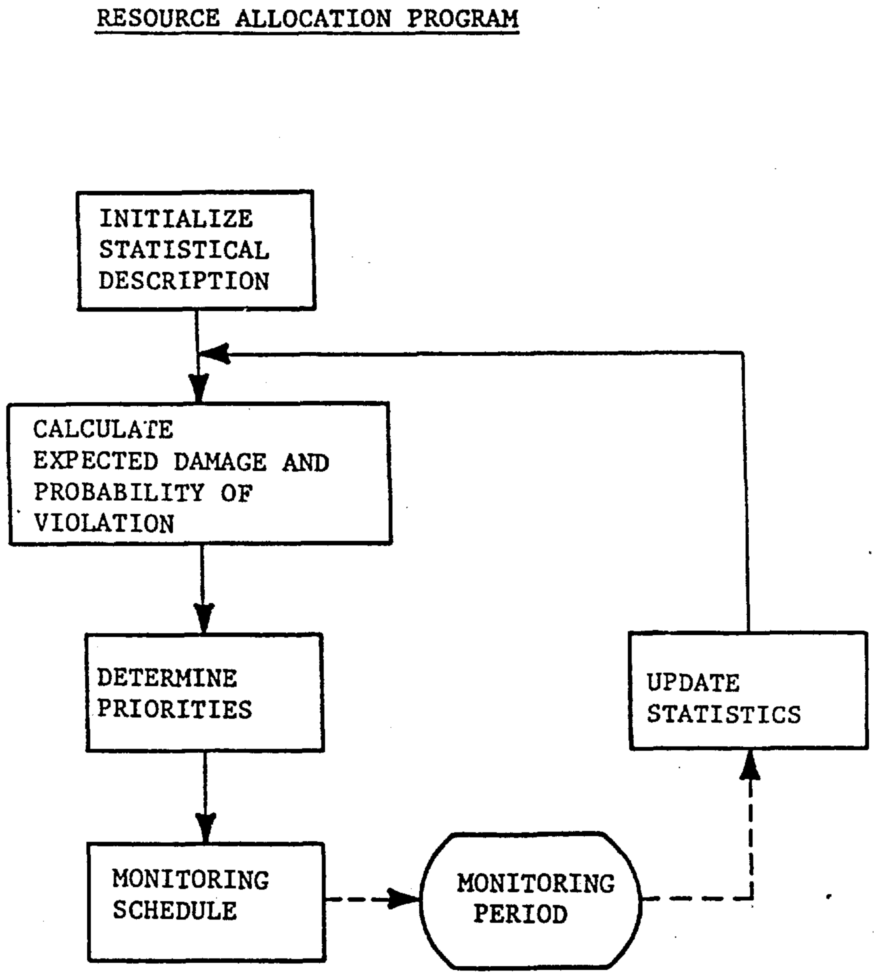

Figure 8.1 Flow of Resource Allocation Program.

(4) Monitoring Schedule

Take the sampling frequencies obtained in the previous component and determine which day to sample which sources.

(5) Monitoring Period

This box represents the actual time spent monitoring the sources.

(6) Update Statistics

Combine new self-monitoring and compliance data with the initial statistics to obtain an updated statistical description of the effluents.

All the components except the "Monitoring Schedule" have been described in detail in Sections V, VI, and VII. The scheduling of the sampling depends on a number of factors which are difficult to quantify in an optimization framework, such as: the spatial location of the various effluent sources, the size of the monitoring agency's jurisdiction, and the availability of personnel. This scheduling is beyond the scope of this report.

Figure 8.2 gives a more detailed description of the Resource Allocation Program. It describes in detail what data are needed by each component of the Program. The basic output of the Program is the priorities and the monitoring frequencies.

VIII.2 SIMPLIFIED EXAMPLE

The performance of the Resource Allocation Program is demonstrated in this section, using a simplified example. Initially, it is assumed that there are four sources to be monitored, each having four months of self-monitoring data available from which to obtain the initial statistics.

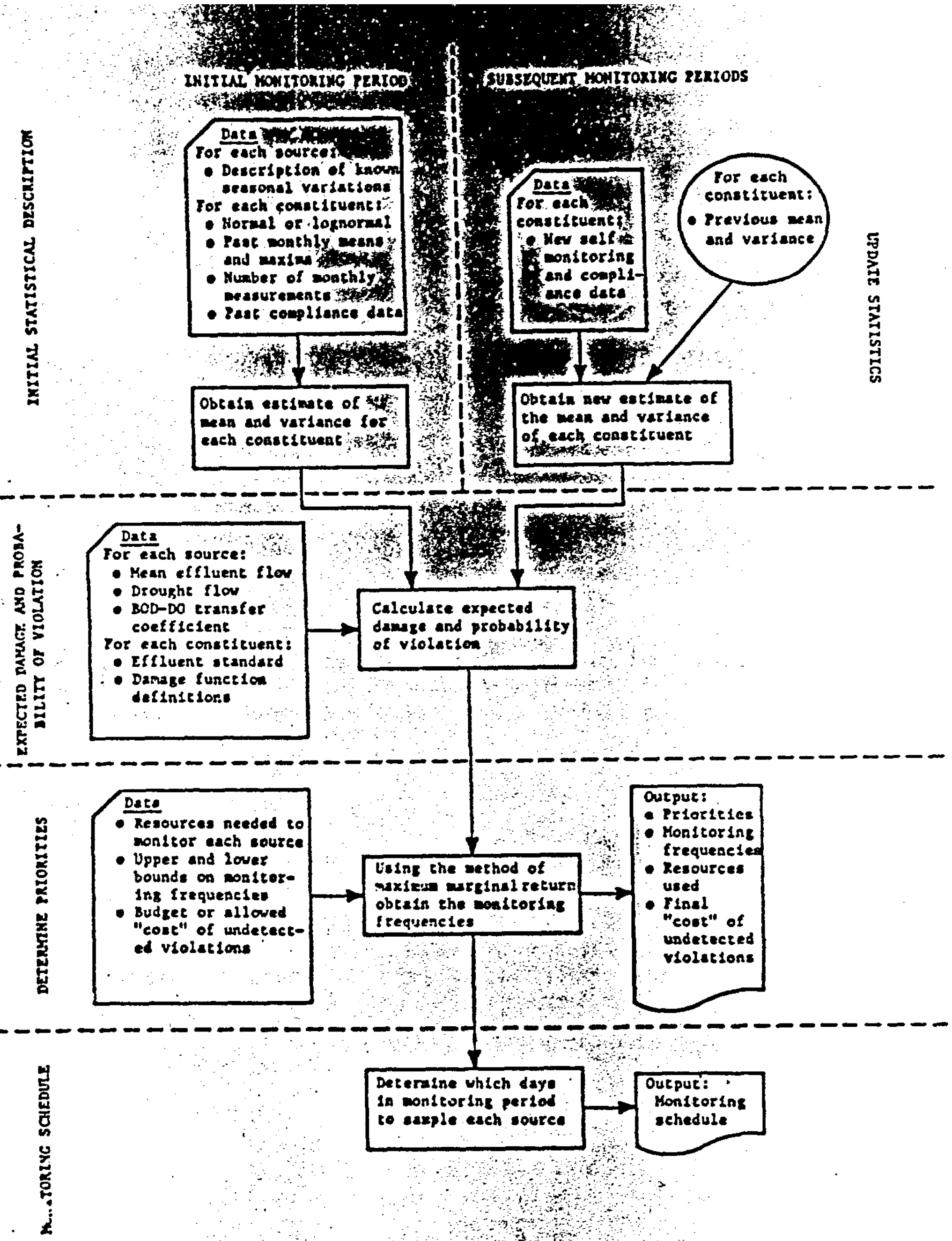

Figure 8.2 Resource allocation program.

The initial self monitoring data assumed are shown in Tables 8.3a through 8.3e. The data have been abstracted from real data that were used for the demonstration case (Section IX). Using the procedure outlined in Section V.3, Tables 8.4a through 8.4e present the initial statistics obtained from the data. The estimated mean and estimated standard deviation are the monthly estimates using the techniques developed in Appendix A. For Source 4, the sample size of the effluent constituents for a single month is 2; therefore, the data in months 1 and 2 and months 3 and 4 have to be aggregated, as discussed in Section V.2. Thus, only two estimates of the mean and two of the variance are given in Table 8.4d and 8.4e. Tables 8.4a through 8.4e also show how the estimates of the mean and standard deviation are sequentially updated as the monthly estimates are combined to obtain the estimates to be used in the Resource Allocation Program. For this case the design parameters k_n and k_ν, which determine the degree of the discounting of past information, have been set to 3.* The updated mean and variance for month 2 are therefore the combined estimates derived from the 1st and 2nd monthly estimates. The updated mean and variance for month 3 are the combination of the updated estimates for month 2 and monthly estimate for month 3. The same process is repeated for month 4, yielding the initial statistical description to be used in the program.

The expected damage and probability of violation obtained from the data are shown in Table 8.5, along with the estimated source flow and the stream flow. For this case, the upstream concentration was assumed to be at a level causing zero damage, and the distributions of the various parameters were assumed uncorrelated. Certain of the entries deserve some comment. Source 3 is a large sewage treatment plant. From the table, the impact of BOD_5 and phosphates is large; however, the standards are also large and therefore the probability of violation for the parameters is small. Source 4 has a relatively small impact on the stream. (i.e., small expected damage); however, the standards have been set so that the probability of violation is very large. The resources required

---

* k_n and k_ν are discussed in Section V.2. The effect of changing k_n and k_ν is shown in VIII.3.

Table 8.3a SELF MONITORING DATA FOR SOURCE 1

Month	Mean source flow, Ml/day	Parameter: pH Max Eff. standard: 9 Distribution: Normal			Parameter: pH Min Eff. standard: 6 Distribution: Normal			Parameter: Lead Eff. standard: 2 kg Distribution: Normal		
		Mean	Max	Sample size	Mean	Min	Sample size	Mean, kg	Max, kg	Sample size
1	0.90	8.5	10.6	20	8.5	6.0	20	0.41	1.0	20
2	1.10	7.6	9.0	22	7.6	5.6	22	1.08	1.7	22
3	1.20	8.3	9.8	22	8.3	6.4	22	1.09	6.3	22
4	0.85	8.1	9.5	20	8.1	6.6	20	0.52	1.8	22

Table 8.3b SELF MONITORING DATA FOR SOURCE 2

Month	Mean source flow, Ml/day	Parameter: Chromium Eff. standard: 0.45 kg Distribution: Normal			Parameter: Copper Eff. standard: 1.5 kg Distribution: Lognormal			Parameter: Fluoride Eff. standard: 30 kg Distribution: Normal		
		Mean, kg	Max, kg	Sample size	Mean, kg	Max, kg	Sample size	Mean kg	Max, kg	Sample size
1	0.80	0.216	0.808	18	0.524	1.89	18	24.4	31.4	18
2	0.78	0.313	0.867	19	0.374	1.87	19	25.4	31.9	19
3	0.87	0.214	0.620	21	0.366	1.25	22	24.7	31.0	22
4	0.85	0.132	0.255	16	0.110	0.42	14	14.0	31.0	11

Table 8.3c SELF MONITORING DATA FOR SOURCE 3

Month	Mean source flow, Ml/day	Parameter: BOD_5 Eff. standard: 3500 kg Distribution: Normal			Parameter: Phosphate Eff. standard: 500 kg Distribution: Lognormal			Parameter: Sus. Solids Eff. standard: 4050 kg Distribution: Lognormal			Parameter: Dissolved oxygen	
		Mean, kg	Max, kg	Sample size	Mean, kg	Max, kg	Sample size	Mean, kg	Max, kg	Sample size	Mean, mg/l	Sample size
1	105	1165	2115	30	178	658	30	2430	6030	30	3.9	30
2	110	900	2115	31	171	338	31	1665	5130	31	3.8	31
3	109	1395	2880	30	171	500	30	3240	10935	30	4.2	30
4	108	1080	2385	31	88	275	31	2160	4590	31	4.1	31

Table 8.3d SELF MONITORING DATA FOR SOURCE 4, PIPE 1

Month	Mean source flow, Ml/day	Parameter: Phosphates Eff. standard: 0.6 kg Distribution: Normal			Parameter: Sus. Solids Eff. standard: 25 kg Distribution: Normal		
		Mean, kg	Max, kg	Sample size	Mean, kg	Max, kg	Sample size
1	0.35	0.15	0.24	2	12.0	18.9	2
2	0.26	0.30	0.36	2	16.6	18.9	2
3	0.29	0.31	0.36	2	16.4	18.0	2
4	0.30	1.20	2.56	2	11.0	15.3	2

Table 8.3e SELF MONITORING DATA FOR SOURCE 4, PIPE 2

Month	Mean source flow, Ml/day	Parameter: Phosphates Eff. standard: 3.5 kg Distribution: Normal			Parameter: Sus. Solids Eff. standard: 80 kg Distribution: Normal		
		Mean, kg	Max, kg	Sample size	Mean, kg	Max, kg	Sample size
1	0.90	2.9	3.2	2	158	296	2
2	1.01	3.5	3.9	2	18	26	2
3	1.09	2.9	3.1	2	93	145	2
4	1.00	5.8	9.8	2	31	33	2